

Linear Invariant Systems Theory for Signal Enhancement

A. M. Tomé , A. R. Teixeira , A. Teixeira, G. Miguel , P. Georgieva , E. W. Lang

Abstract – This paper discusses a linear time invariant (LTI) systems approach to signal enhancement via projective subspace techniques. It provides closed form expressions for the frequency response of data adaptive finite impulse response eigenfilters. An illustrative example using speech enhancement is also presented.

Resumo – Este artigo apresenta a aplicação da teoria de sistemas lineares invariantes no tempo (LTI) na análise de técnicas de sub-espço. A resposta em frequência dos filtros resultantes da decomposição em valores singulares é obtida aplicando as propriedades dos sistemas LTI.

Keywords – Signal Enhancement, SVD, Subspace-Based Methods

I. INTRODUCTION

Signal enhancement via projective subspace techniques is widely used in speech processing and biomedical signal processing to improve noise-corrupted signals. Different notations or mathematical formalisms are discussed in the literature [1], [2], [3], but singular value decomposition (SVD) is common to all. A data-derived trajectory matrix \mathbf{X} is replaced by a low-rank approximation $\hat{\mathbf{X}}$. From this low-rank matrix $\hat{\mathbf{X}}$ an enhanced version $\hat{x}[k]$ of the original signal is then obtained employing diagonal averaging [4]. A discussion of these techniques in the frequency domain was given by [5] where it was shown that signal enhancement can be achieved by a bank of finite impulse response filters arranged as parallel pairs of analysis-synthesis filters. The proposed approach was formulated using matrix algebra operations [5], [6], [1]. In this work we show that linear invariant system theory provides alternative tools to derive such input-output relations. Analytical expressions for the frequency response of these filters will be given. In addition, the main filter characteristics, i.e. causality and being zero-phase, are deduced. The paper addresses the following three topics: the meaning of SVD in terms of univariate time series analysis; the filter bank interpretation and a few illustrative examples using speech signals.

II. UNIVARIATE TIME-SERIES AND SVD

Singular value decomposition constitutes the main tool to estimate subspace models for multidimensional data sets. In multi-sensor signal processing, the data vector is naturally formed with samples of different sensors. However, projective subspace techniques can also be applied to univariate time series by forming vectors with sliding windows of the signal. This transformation is called embedding the signal into the space of time-delayed coordinates. Considering a segment of a signal $(x[1], x[2], \dots, x[K])$, the multidimensional signal is obtained by $\mathbf{x}_{k*} = (x[k], \dots, x[k +$

$N])$, $k = 1, \dots, M = K - N + 1$. The lagged vectors lie in a space of dimension N and constitute the rows of the trajectory matrix

$$\mathbf{X} = \begin{bmatrix} x[1] & x[2] & \dots & x[N] \\ x[2] & x[3] & \dots & x[N+1] \\ x[3] & x[4] & \dots & x[N+2] \\ \vdots & \vdots & \ddots & \vdots \\ x[M] & x[M+1] & \dots & x[K] \end{bmatrix} \quad (1)$$

Note that this matrix has identical entries along its anti-diagonals, hence forms an Hankel matrix [1]. There are other alternatives to form the data matrix via embedding the signal in an N -dimensional space which yield a Toeplitz matrix. The latter has identical elements along its diagonals [2]. However, the processing steps are the same, only adapted to cope with the differences in data organization [7]. Following that strategy, the univariate signal is organized into an $M \times N$ matrix \mathbf{X} whose SVD [8] allows to explain the data set as a product of matrices $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are orthonormal matrices with dimension $M \times M$ and $N \times N$, respectively. The matrix $\mathbf{\Sigma}$ is an $M \times N$ matrix with $r \leq \min(M, N)$ non-zero singular values along the diagonal and zeros everywhere else. The eigenvector matrices (\mathbf{U} and \mathbf{V}) result from an eigenvalue decomposition of two symmetric and square matrices computed from the data matrix. Assuming that the vectors \mathbf{x}_{k*} , $k = 1, 2, \dots, M$, represent the rows of \mathbf{X} , the two different square matrices are computed in the following way:

- Matrix $\mathbf{S} = \mathbf{X}^T \mathbf{X}$ is an $N \times N$ matrix where each entry represents the correlations between pairs of entries of the data vectors. It is an outer product matrix corresponding to the non-normalized correlation matrix. If the data is centered, it is also the non-normalized covariance matrix, also called scatter matrix. Its eigenvalue decomposition reads

$$\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}^T.$$

The matrix $\mathbf{\Lambda}$ is a diagonal matrix with at most $r \leq \min(M, N)$ non-zero eigenvalues corresponding to the square of the singular values. And the eigenvector matrix \mathbf{V} is orthonormal, e.g. $\mathbf{V}^T \mathbf{V} = \mathbf{I}$, where \mathbf{I} is the identity matrix.

- Matrix $\mathbf{K} = \mathbf{X}\mathbf{X}^T$ is an $M \times M$ matrix where each entry represents the dot product between pairs of vectors of the data set (the rows of \mathbf{X}). It is known as kernel matrix or dot product matrix. The eigenvalue decomposition of this matrix reads

$$\mathbf{K} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T.$$

The matrix $\mathbf{\Lambda}$ is a diagonal matrix with at most $r \leq$

$\min(M, N)$ non-zero eigenvalues and $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ is an orthonormal eigenvector matrix.

Orthogonal subspace models, like SVD or principal component analysis (PCA), are described solely by the matrix \mathbf{V} that defines an orthonormal basis vector matrix of the $N - \text{dimensional}$ space of the data [8]. A low-rank approximation of the data matrix \mathbf{X} can be expressed as follows

$$\hat{\mathbf{X}} = \mathbf{XVPV}^T = \mathbf{YPV}^T \quad (2)$$

where each term means:

- matrix $\mathbf{Y} = \mathbf{XV}$ represents the projection of the data vectors $\mathbf{x}_{k*} \in \mathbb{R}^N$ onto the basis vectors. Each column \mathbf{y}_{*n} represents the projections of all row vectors \mathbf{x}_{k*} , $k = 1, \dots, M$ onto the n -th basis vector \mathbf{v}_{*n} .
- matrix \mathbf{P} is a diagonal matrix with diagonal entries equal to $0 \leq p_{nn} \leq 1$. If $p_{nn} = 1$, the $n - \text{th}$ column of \mathbf{Y} is retained, and if $p_{nn} = 0$, it is replaced by a null vector. In the general case, the scaling factors p_{nn} can be estimated from the eigenvalue spectrum and the related noise variance [1].
- the reconstructed data $\hat{\mathbf{X}}$ in the original $N - \text{dimensional}$ space results from the product with \mathbf{V}^T

Thus the reconstructed data matrix $\hat{\mathbf{X}}$ corresponds to a reduced rank approximation of the original data matrix with a possibly modified eigenvalue spectrum. But the reconstruction does not preserve the Hankel structure of the original data matrix (see eqn. 1). To rectify this, each anti-diagonal element is substituted by the corresponding average of all entries of the anti-diagonal. Finally, the embedding is reversed to obtain the reconstructed signal $\hat{x}[k]$ which is an enhanced version of the original signal.

III. SVD AS FILTER BANKS

Signal enhancement as it was sketched above can also be addressed employing linear invariant systems theory. In the following we discuss the application of a bank of finite impulse response (FIR) filters, where analysis and synthesis filter pairs are connected in parallel. Note that eqn. (2) can be expressed as a weighted superposition of terms related with the non-zero singular values/eigenvalues and corresponding eigenvectors of the subspace model. Employing the subspace model $\mathbf{V} = [\mathbf{v}_{*1}, \mathbf{v}_{*2}, \dots, \mathbf{v}_{*N}]$, and using block matrix operations, the terms $\hat{\mathbf{X}}_n$ are $M \times N$ matrices of rank one. The reconstructed data matrix can then be expressed as

$$\begin{aligned} \hat{\mathbf{X}} &= \mathbf{Xv}_{*1}p_{11}\mathbf{v}_{*1}^T + \dots + \mathbf{Xv}_{*N}p_{NN}\mathbf{v}_{*N}^T \\ &= \mathbf{y}_{*1}p_{11}\mathbf{v}_{*1}^T + \dots + \mathbf{y}_{*N}p_{NN}\mathbf{v}_{*N}^T = \sum_{n=1}^N \hat{\mathbf{X}}_n(3) \end{aligned}$$

where \mathbf{v}_{*n} , $n = 1, \dots, N$ represent analysis or synthesis filter coefficients. In Hansen *et al.* [5], a filter bank architecture is proposed where formally analysis and synthesis

filters operate on trajectory matrices with a Hankel and a Toeplitz structure, respectively, to include the diagonal averaging [4] during synthesis.

However, in the framework of linear invariant systems theory, the filter bank structure needed to achieve the output time series $\hat{x}[k]$ should be provided by the input time series $x[k]$ instead of by the trajectory matrix. Hence we propose an approach based on filter responses and related transfer functions rather than on matrix manipulations.

As mentioned above, each column \mathbf{y}_{*n} , $n = 1, \dots, N$ of the projected data \mathbf{Y} is obtained via $\mathbf{y}_{*n} = \mathbf{Xv}_{*n}$. Each element of the $M \times 1$ vector \mathbf{y}_{*n} is the dot product between the $n - \text{th}$ eigenvector and a row of the data matrix. But this data manipulations can also be formulated as the weighted sum of a sequence of samples of the original time series,

$$y_n[k] = \sum_{i=1}^N v_{in}x[k+i-1] \quad (4)$$

where $1 \leq k < M$ and $y_n[k]$ are the elements of the n th column of matrix \mathbf{Y} , i.e. \mathbf{y}_{*n} . Therefore, the \mathbf{y}_{*n} has M samples starting by time index $k = 1$, like in the first column of the matrix \mathbf{X} . The entries of the vector \mathbf{v}_{*n} , the $n - \text{th}$ column of the subspace model, are the coefficients of an anti-causal finite impulse response (FIR) filter as the output at time index k depends on input samples at time index $k, k+1 \dots k+N-1$.

The transfer function $H_n(z)$ of the analysis step can be computed by substituting in eqn. (4) every delay operation by the corresponding z transform [9]. Therefore by transforming $x[k]$ to $X(z = \sum_{-\infty}^{\infty} x[k]z^{-k})$, $x[k \pm d]$ to $z^{\pm d}X(z)$ and $y_n[k]$ to $Y_n(z)$, the filtering operation can be formulated using the following transfer function

$$H_n(z) = \frac{Y_n(z)}{X(z)} = (v_{1n} + v_{2n}z^1 + \dots + v_{Nn}z^{(N-1)}) \quad (5)$$

This transfer function $H_n(z)$, $n = 1, \dots, N$ results from an output-input ratio and constitutes the analysis block as it decomposes the input into several components $y_n[k]$, $n = 1, \dots, N$. In filter bank terminology, the analysis filter is then followed by the synthesis filter which should combine components to form a new signal. To facilitate the exposition, let's consider the n -th term of eqn. (3) and assume $p_{nn} = 1$ for simplicity. In that case the n -th contribution to the reconstructed data matrix is

$$\hat{\mathbf{X}}_n = \mathbf{Xv}_{*n}\mathbf{v}_{*n}^T = \mathbf{y}_{*n}\mathbf{v}_{*n}^T. \quad (6)$$

Therefore each column of the rank-one matrix $\hat{\mathbf{X}}_n$ is a scaled version of \mathbf{y}_{*n}

$$\hat{\mathbf{X}}_n = \begin{bmatrix} v_{1n}y_n[1] & v_{2n}y_n[1] & \dots & v_{Nn}y_n[1] \\ v_{1n}y_n[2] & v_{2n}y_n[2] & \dots & v_{Nn}y_n[2] \\ v_{1n}y_n[3] & v_{2n}y_n[3] & \dots & v_{Nn}y_n[3] \\ \vdots & \vdots & \dots & \vdots \\ v_{1n}y_n[M] & v_{2n}y_n[M] & \dots & v_{Nn}y_n[M] \end{bmatrix} \quad (7)$$

Obviously, the resulting matrix does not have the Hankel structure of the original matrix \mathbf{X} . But by replacing the entries in each anti-diagonal of $\hat{\mathbf{X}}_n$ by their average, an Hankel matrix is obtained again. Interestingly, the diagonal averaging can equally well be formulated as a linear filtering operation

$$\hat{x}_n[k] = \frac{1}{N_d} \sum_{i=l}^s v_{in} y_n[k - i + 1] \quad (8)$$

where the quantities N_d , l and s have values according to number of elements in the anti-diagonals of the matrix defined in eqn. (7). More specifically, the response can be sub-divided into a transient and a steady state response according to the following distinction:

- With N elements, eqn. (8) represents a steady state response of the filter in the case of $N \leq k \leq M$ and we have $N_d = N$, $l = 1$ $s = N$.
- With less than N elements, eqn. (8) represents the transitory response of the filter:
 - if $1 \leq k \leq (N - 1)$ (upper left corner of the matrix $\hat{\mathbf{X}}_n$) then we have $N_d = k$, $l = 1$ and $s = N_d$;
 - if $(M+1) \leq k \leq K$ (lower right corner of the matrix $\hat{\mathbf{X}}_n$) then we have $N_d = K - k + 1$, $l = N - N_d + 1$ and $s = N$.

Note that the entries of the vector \mathbf{v}_{*n} , the n -th column of the subspace model, are the coefficients of a causal finite impulse response (FIR) filter. Both cases can be unified by formally setting $y_m[k] = 0$ for the time indices $k < 1$ and $k > M$, and always compute eqn. (8) as in the steady-state case. Therefore, the transfer function for the synthesis filter reads

$$F_n(z) = \frac{\hat{X}_n(z)}{Y_n(z)} = \frac{1}{N} (v_{1n} + v_{2n}z^{-1} + \dots + v_{Nn}z^{-(N-1)}) \quad (9)$$

Notice that the analysis ($H_n(z)$) and synthesis ($F_n(z)$) transfer functions differ by a scale factor ($1/N$) and by the sign of the powers of z . Therefore the magnitudes of the frequency response of both filters are related by a scale factor ($1/N$) and their phases are symmetric. The transfer function of the global system is a cascade formed by the projection step (analysis) and the reconstruction step with diagonal averaging (synthesis) given by

$$T_n(z) = \frac{\hat{X}_n(z)}{X(z)} = F_n(z)H_n(z) = \sum_{i=-(N-1)}^{N-1} t_{in}z^i \quad (10)$$

The coefficients t_{in} result from the product of two polynomials with the same coefficients but powers of z with opposite sign, and then $t_{in} = t_{-in}$, $i = 1, \dots, (N - 1)$ [9]. Therefore, the frequency response $T_n(e^{j\omega})$ has the following expression

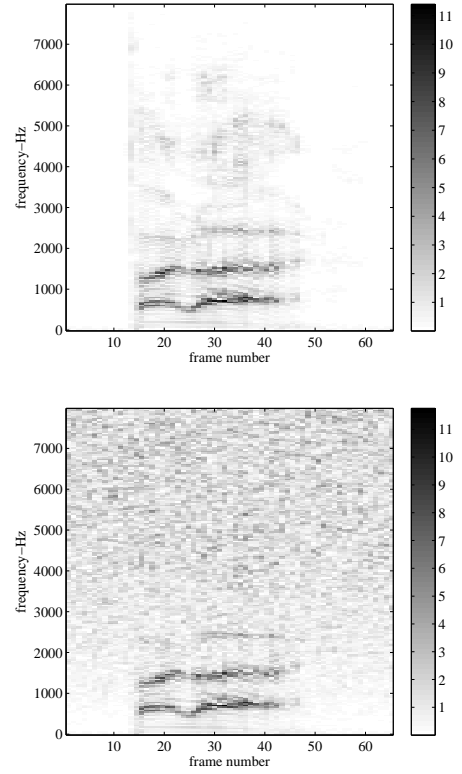


Fig. 1 - Spectrogram of the signal with different SNR: *left*-SNR=40dB, *right*-SNR=10dB.

$$T_n(e^{j\omega}) = t_{0n} + \sum_{i=1}^{N-1} 2t_{in} \cos(i\omega) \quad (11)$$

where $j = \sqrt{-1}$. The frequency response is a periodic real function, with period equal $\omega = 2\pi$, hence corresponds to a zero-phase filter. This means that each extracted component $\hat{x}_n[k]$ is always in-phase with its related original $x[k]$. The orthogonal output sequences $y_n[k]$, $n = 1 \dots N$ of the analysis step are filtered versions of the input sequence $x[k]$ and their energy content is given by the eigenvalue associated with the corresponding eigenvector (eigenfilter). The scale factor p_{nn} only changes the amplitude of the sequence $y_n[k]$. The total transfer function is finally obtained by adding the transfer functions of the parallel branches of the filter bank. The resulting output $\hat{x}[k]$ is a sum of the selected signals $\hat{x}_n[k]$, e.g, the outputs of the cascaded filter pairs formed by $H_n(z)$ and $F_n(z)$. Notice, that the embedding of the time-series as suggested by (1) leads to an *anti-causal* filter for the analysis step and to a *causal* filter for the synthesis step. Using alternative embedding procedures, this property of the filters can interchange. Though examples of frequency responses of the eigenfilters are shown graphically in [5], no analytical expressions of the filter responses are given. Instead, the present work also deduces such closed-form analytical expressions for the analysis and synthesis filters. Note that frequency responses of the component filters $H_n(z)$ and $F_n(z)$ cannot be given in closed-form similar to $T_n(z)$ in (11) due to lacking symmetry properties of their coefficients [9]. But notice that the absolute values of the frequency responses of all the filters have the

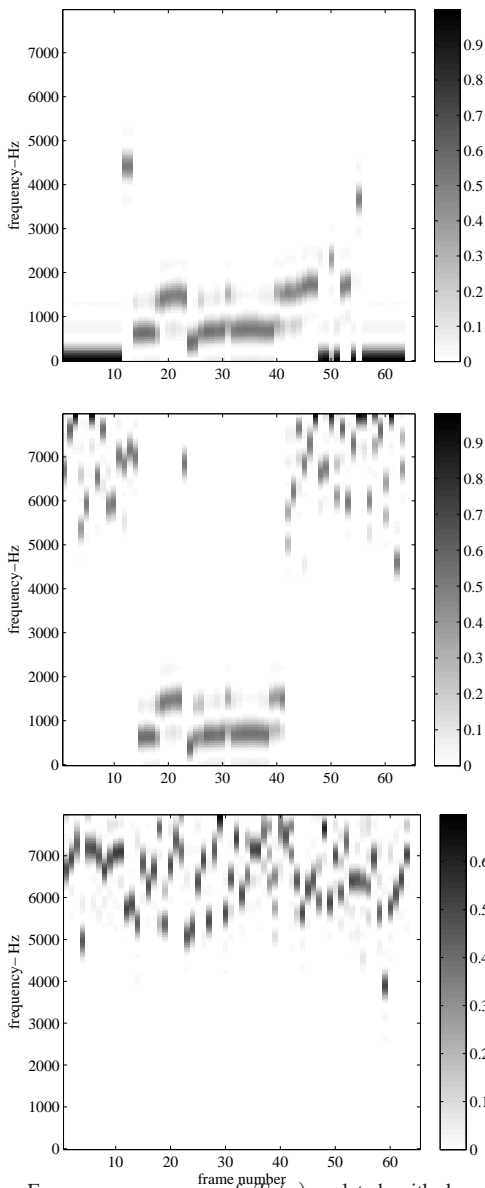


Fig. 2 - Frequency responses of $T_1(z)$, related with largest eigenvalue, along the frames computing model with :*top*-SNR=40dB, *middle*-SNR=10dB, *bottom*-SNR=0dB

same profile.

IV. DISCUSSION AND CONCLUSION

In this section we present illustrative examples of the frequency response of eigenfilters for a segment of a speech signal corrupted by additive gaussian noise.

Fig. 1 shows the frequency content of a noisy speech signal in each frame (spectrogram). Each frame correspond to a segment of 25ms (400 samples) with 60% of overlap between the frames. In the first spectrogram the energy of the signal is concentrated in the frames 15 – 45 corresponding to frequencies $f < 2000Hz$. With decreasing the SNR, the speech signal becomes increasingly corrupted and the frequency content is distributed more uniformly (see Fig. 1, right). The orthogonal subspace model \mathbf{V} was computed for each frame using $K = 400$ samples and an embedding dimension $N = 30$.

Fig 2 shows the frequency response of $T_1(z)$, i.e. of

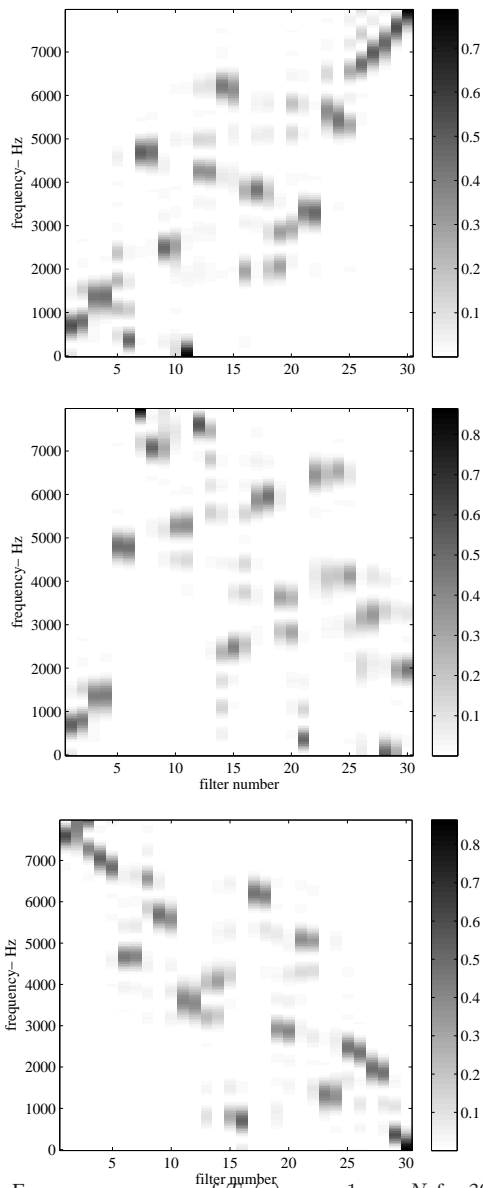


Fig. 3 - Frequency responses of $T_n(z), n = 1, \dots, N$ for 30th frame using:*top*-SNR=40dB, *middle*-SNR=10dB, *bottom*-SNR=0dB

the analysis/synthesis filter that corresponds to the largest eigenvalue of the subspace models. This filter is clearly centered in the region of the highest energy of the input signals. If the SNR is high (figure on left), the passband of the filter matches the signal. Even when the energy of the signal is low, the filter has its passband in low frequency range. When the noise level increases, the passband of the filter moves towards the high frequency range for the frames without signal information (see figure on middle). But when $SNR = 0dB$, the filter is always centered on the high-frequency range both in frames without or with active voice signals. Fig. 3 shows the frequency response of $T_n(z), n = 1 \dots 30$ computed using a frame with active voice (the 30th frame). The differences on the profile of the frequency responses is obvious. The filters were ordered according the values of the eigenvalues. It can be seen that when the SNR is high, the first filters, associated with the largest eigenvalues, have their passband centered in the low frequency range, and the last filters, as-

sociated with the smallest eigenvalues, are centered in the high-frequency range. Decreasing the SNR, the filters associated with the highest eigenvalues have passbands both in the low-frequency band and in the high-frequency band (see figure on middle). However, for $SNR = 10dB$, the first five filters still concentrate their passband in low frequency range. Finally, when $SNR = 0dB$, the filters with their passband in the high frequency range are the ones corresponding to the largest eigenvalues, while the filters centered in low frequency range correspond to the smallest eigenvalues.

The interpretation of subspace-based methods as filter banks helps to attain a clear-cut insight into the outcomes of the method. By applying a linear invariant system theory approach, analytical expressions of the frequency response are deduced in this work. These results thus corroborate the properties of the SVD steps referred to in previous works [5], [4], [6].

By the frequency responses of the filter bank, corresponding to the basis vectors of the subspace model, the frequency content of the different components can be attained easily. Eigenfilters are data adaptive, and the relevance of one component to the energy of the input signal is deduced from the corresponding eigenvalue. Moreover, the frequency profile of each component is determined only at the projection step. However, in order to get a component in phase with the input signal, the diagonal averaging is required.

REFERENCES

- [1] Per Christian Hansen and S. Holdt Jensen, "Subspace-based noise reduction for speech signals via diagonal and triangular matrix decompositions: Survey and analysis", *Eurasip Journal on Advances in Signal Processing*, vol. Vol 2007, 2007.
- [2] Nina Golyandina, Vladimir Nekrutkin, and Anatoly Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*, Chapman & HALL/CRC, 2001.
- [3] M. Ghil, M.R. Allen, M. D. Dettinger, and K. Ide, "Advanced spectral methods for climatic time series", *Reviews of Geophysics*, vol. 40, no. 1, pp. 3.1–3.41, 2002.
- [4] I. Dologlou and G. Carayannis, "Physical interpretation of signal reconstruction from reduced rank matrices", *IEEE Transactions on Signal Processing*, vol. 39, no. 7, pp. 1681–1682, 1991.
- [5] Per Christian Hansen and Soren Holdt Jensen, "FIR filter representations of reduced-rank noise reduction", *IEEE Transactions on Signal Processing*, vol. 46, no. 6, pp. 1737–1741, 1998.
- [6] Kris Hermus, Patrick Wambacq, and Hugo Van hamme, "A review of signal subspace speech enhancement and its application to noise robust speech recognition", *Eurasip Journal on Advances in Signal Processing*, 2007.
- [7] A. R. Teixeira, A. M. Tomé, M. Böhm, C. G. Puntonet, and E. W. Lang, "How to apply non-linear subspace techniques to univariate biomedical time series", *IEEE Transactions on Instrumentation & Measurement*, vol. 58, no. 8, pp. 2433–2443, 2009.
- [8] K. I. Diamantaras and S.Y. Kung, *Principal Component Neural Networks, Theory and Applications*, Wiley, 1996.
- [9] Leland B. Jackson, *Digital Filters and Signal Processing*, Kluwer Academic Publishers, 1996.